



CGCTCAGGATAGACTTCCGCGTAGAGATCGGATCCCGGCGGATTATTATATACCTCGATCGATC
 TTCTCTATATACCGCGGATATGGGATATATACACACACACCGCGCGGATAGCATGACTGATCTA
 CCCCACCTTCCGCGGATATGGGATATATACACACACACCGCGCGGATAGCATGACTGATCTA
 CACAGACTTACGCTTCTCACTTACTTACTAACCAATTCGGGAGAGGGGCGGATTCGGCGGAGG

PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

OMIM

Boo

Search for

Limits

Preview/Index

History

Clipboard

Details

Show:

☐ 1: AJ005830. Escherichia coli ...[gi:3123495]

Links

LOCUS EC05830 2458 bp DNA linear BCT 08-MAR-2000

DEFINITION Escherichia coli tatABCD operon.

ACCESSION AJ005830

VERSION AJ005830.1 GI:3123495

KEYWORDS mttB gene; mttC gene; Sec-independent protein translocase; tata gene; TataA protein; tatB gene; TatB protein; tatC gene; TatC protein; tatD gene; TatD protein; yigU gene; yigW gene.

SOURCE Escherichia coli

ORGANISM Escherichia coli

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.

REFERENCE 1

AUTHORS Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y.

TITLE The complete genome sequence of Escherichia coli K-12

JOURNAL Science 277 (5331), 1453-1474 (1997)

MEDLINE 97426617

PUBMED 9278503

REFERENCE 2

AUTHORS Sargent, F., Bogsch, E.G., Stanley, N.R., Wexler, M., Robinson, C., Berks, B.C. and Palmer, T.

TITLE Overlapping functions of components of a bacterial Sec-independent protein export pathway

JOURNAL EMBO J. 17 (13), 3640-3650 (1998)

MEDLINE 98315056

PUBMED 9649434

REFERENCE 3 (bases 1 to 2458)

AUTHORS Palmer, T.

TITLE Direct Submission

JOURNAL Submitted (30-APR-1998) Palmer T., Nitrogen Fixation Laboratory, John Innes Centre, NORWICH, NR4 7UH, UNITED KINGDOM

COMMENT Related sequence: AE000459.

FEATURES

source

Location/Qualifiers

1..2458

/organism="Escherichia coli"

/mol_type="genomic DNA"

/strain="K-12"

/sub_strain="MC4100"

/db_xref="taxon:562"

gene

56..325

/gene="tataA"

CDS

56..325

/gene="tataA"

/function="Sec-independent protein translocase"

/codon_start=1

/transl_table=11
/product="TatA protein"
/protein_id="CAA06724.1"
/db_xref="GI:3123496"
/db_xref="SWISS-PROT:O65938"
/translation="MGGISIWQLLIIAVIVVLLFGTKKLGSIGSDLGASIKGFKKAMS
DDEPKQDKTSQDADFTAKTIADKQADTNQEQAKTEDAKRHDKEQV"
gene 329..844
/gene="tatB"
CDS 329..844
/gene="tatB"
/function="Sec-independent protein translocase"
/codon_start=1
/transl_table=11
/product="TatB protein"
/protein_id="CAA06725.1"
/db_xref="GI:3123497"
/db_xref="SWISS-PROT:O69415"
/translation="MFDIGFSELLLVFIIGLVVLGPQRLPVAVKTVAGWIRALRSLAT
TVQNELTQELKLQEFQDSLKKVEKASLTNLTPELKASMDLRQAESMKRSYVANDPE
KASDEAHTIHNPPVVKDNEAAHEGVTPAAAQTQASSPEQKPETTPEPVVKPAADAEPT
AAPSPSSSDKP"
conflict 380..384
/gene="tatB"
/citation=[1]
/replace="gccgcct"
gene 847..1623
/gene="tatC"
CDS 847..1623
/gene="tatC"
/function="Sec-independent protein translocase"
/note="alternative gene names: yigU, mttB"
/codon_start=1
/transl_table=11
/product="TatC protein"
/protein_id="CAA06726.1"
/db_xref="GI:3123498"
/db_xref="SWISS-PROT:P27857"
/translation="MSVEDTQPLITHLIELRKRLLNCCIIVIVIFLCLVYFANDIYHL
VSAPLIKQLPQGSTMATDVASPFFTPIKLTFMVSLILSAPVILYQVWAFIAPALYKH
ERRLVVPLLVSLLLFYIGMAFAYFVVFPPLAFGFLANTAPEGVQVSTDIASYSFVMA
LFMAFGVSFEVPVAIVLLCWMGITSPEDLRKKRPYVLVGAFVGMMLTTPPDVFSQTL
AIPMYCLFEIGVFFSRFVVGKGRNREEENDAEAESEKTEE"
gene 1653..2447
/gene="tatD"
CDS 1653..2447
/gene="tatD"
/function="Sec-independent protein translocase"
/note="alternative gene names: yigW, mttC"
/codon_start=1
/transl_table=11
/product="TatD protein"
/protein_id="CAA06727.1"
/db_xref="GI:3123499"
/db_xref="SWISS-PROT:P27859"
/translation="MEYRMFDIGVNLTSQFAKDRDDVVACAFDAGVNGLLITGTNLR
ESQQAQKLARQYSSCWSTAGVHPHDSSQWQAATEEAIIELAQPEVVAIGECGLDFNR
NFSTPEEQERAFVAQLRIAADLNMPVFMHCRDAHERFMTLLEPWLDKLPAGAVLHCFTG
TREEMQACVAHGIYIGITGWVCDERRGLELRELLPLIPAELLIIETDAPYLLPRDLTP
KPSSRRNEPAHLPHILQRIAHWRGEDAAWLAATTDANVKTLFGIAF"

conflict

2204..2205

/gene="tatD"

/citation=[1]

/replace="acc"

BASE COUNT 573 a 604 c 684 g 597 t
ORIGIN

```
1  cgtgtaacgt  ataatgcggc  tttgtttaat  catcatctac  cacagaggaa  catgtatggg
61  tggatcacgt  atttggcagt  tattgattat  tgccgtcatc  gttgtactgc  tttttggcac
121  caaaaagctc  ggctccatcg  gttccgatct  tgggtgcgtc  atcaaaggct  ttaaaaaagc
181  aatgagcgat  gatgaaccaa  agcaggataa  aaccagtcag  gatgctgatt  ttactgcgaa
241  aactatcgcc  gataagcagg  cggatacgaa  tcaggaaacag  gctaaaacag  aagacgcgaa
301  gcgccacgat  aaagagcagg  tgtaatccgt  gtttgatata  ggtttttagc  aactgctatt
361  ggtgttcac  atcggcctcg  tcgttctggg  gccgcaacga  ctgcctgtgg  cggtaaaaac
421  ggtagcgggc  tggattcgcg  cgttgcgttc  actggcgaca  acggtgcaga  acgaactgac
481  ccaggagtta  aaactccagg  agtttcagga  cagtctgaaa  aaggttgaaa  aggcgagcct
541  cactaacctg  acgcccgaac  tgaaagcgtc  gatggatgaa  ctacgccagg  ccgcggagtc
601  gatgaagcgt  tcctacgttg  caaacgatcc  tgaaaaggcg  agcgatgaag  cgcacaccat
661  ccataacccg  gtggtgaaa  ataataaagc  tgcgcagtgc  ggcgtaacgc  ctgccgctgc
721  acaaacgcag  gccagttcgc  cggaaacagaa  gccagaaacc  acgccagagc  cggtggtaaa
781  acctgctgcg  gacgctgaac  cgaaaaccgc  tgcaccttc  cttcgtcga  gtgataaacc
841  gtaaacaatg  ctgtagaaga  tactcaaccg  cttatcacgc  atctgattga  gctgcgtaag
901  cgtctgctga  actgcattat  cgcggtgatc  gtgatattcc  tgtgtctggt  ctatttcgcc
961  aatgacatct  atcacctggt  atccgcgcca  ttgatcaagc  agttgccgca  aggttcaacg
1021  atgatcgcca  ccgacgtggc  ctgcgcgttc  tttacgccga  tcaagctgac  ctttatgggtg
1081  tcgctgattc  tgtcagcgcc  ggtgattctc  tatcaggtgt  gggcatttat  cgcccagcg
1141  ctgtataaag  atgaacgtcg  cctggtgggt  ccgctgctgg  tttccagctc  tctgctgttt
1201  tatatcgcca  tggcattcgc  ctactttgtg  gtctttccgc  tggcatttgg  cttccttgcc
1261  aataccgcgc  cggaaagggg  gcaggatatc  accgacatcg  ccagctattt  aagcttcggt
1321  atggcgctgt  ttatggcggt  tgggtgtctc  tttgaagtgc  cggtagcaat  tgtgctgctg
1381  tgctggatgg  ggattacctc  gccagaagac  ttacgcaaaa  aacgcccgtg  tgtgctgggt
1441  ggtgcattcg  ttgtcgggat  gttgctgacg  ccgcccgatg  tcttctcgca  aacgctgttg
1501  gcgatcccga  tgtactgtct  gtttgaaatc  ggtgtcttct  tctcacgctt  ttacgttggt
1561  aaagggcgaa  atcgggaaga  ggaaaacgac  gctgaagcag  aaagcgaaaa  aactgaagaa
1621  taaattcaac  cgcccgtcag  ggcggttgct  atatggagta  caggatgttt  gatatcggcg
1681  ttaatttgac  cagttcgcaa  tttgcgaaag  accgtgatga  tgttgtagcg  tgcgcttttg
1741  acgcgggagt  taatgggcta  ctcatcacgc  gcactaacct  gcgtgaaagc  cagcaggcgc
1801  aaaagctggc  gcgtcagtat  tcgtcctgtt  ggtcaacggc  gggcgtagat  cctcacgaca
1861  gcagccagtg  gcaagctgcg  actgaagaag  cgattattga  gctggccgcg  cagccagaag
1921  tgggtggcgat  tggatgaatg  ggtctcgact  ttaaccgcaa  cttttcgacg  ccggaagagc
1981  aggaacgcgc  ttttggtgac  cagctacgca  ttgccgcaga  ttttaacatg  ccggtattta
2041  tgcactgtcg  cgatgcccac  gagcggttta  tgacattgct  ggagccgtgg  ctggataaac
2101  tgccctgggt  gggtcttcat  tgctttaccg  gcacacgcga  agagatgcag  gcgtgcgtgg
2161  cgcattggaat  ttatatcggc  attaccggtt  gggtttgca  tgaacgacgc  ggactggagc
2221  tgcgggaact  tttgccgttg  attccggcgg  aaaaattact  gatcgaaact  gatgcgccgt
2281  atctgctccc  tcgcgatctc  acgcccgaag  catcatcccg  gcgcaacgag  ccagcccctc
2341  tgccccatat  tttgcaacgt  attgcgcact  ggcgtggaga  agatgccgca  tggctggctg
2401  ccaccacgga  tgctaattgc  aaaacactgt  ttgggattgc  gtttttagag  ttgcggaa
```

//

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

May 20 2003 11:20:12

have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (<http://cgsc.biology.yale.edu>). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (<http://www.genetics.wisc.edu>). *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.

FEATURES	Location/Qualifiers
<u>source</u>	1..11057 /organism="Escherichia coli K12" /mol_type="genomic DNA" /strain="K12" /sub_strain="MG1655" /db_xref="taxon:83333"
<u>promoter</u>	complement(8..36) /note="factor Sigma70; predicted +1 start at 4013851"
<u>promoter</u>	27..56 /note="factor Sigma70; predicted +1 start at 4013913"
<u>gene</u>	168..929 /gene="udp" /note="synonym: b3831"
<u>CDS</u>	168..929 /gene="udp" /EC_number="2.4.2.3" /function="enzyme; Salvage of nucleosides and nucleotides" /note="o253; 100 pct identical to UDP_ECOLI SW: P12758" /codon_start=1 /transl_table=11 /product="uridine phosphorylase" /protein_id="AAC76834.1" /db_xref="GI:1790265" /translation="MSKSDVFHLGLTKNDLQGATLAIVPGDPDRVEKIAALMDKPVKL ASHREFTTWRAELDGKPVIVCSTGIGGPSTSI AVEELAQLGIRTFRLRIGTTGAIQPHI NVGDVLVTTASVRLDGASLHFAPLEFPVADFECTTALVEAAKSIGATTHVGV TASSD TFYPGQERYD TYSGRVVRHFKGSMEEWQAMGVMNYEMESATLLTMCASQGLRAGMVAG VIVNRTQQEIPNAETMKQTESHAVKIVVEAARRLL"
<u>promoter</u>	1009..1037 /note="factor Sigma70; predicted +1 start at 4014894"
<u>gene</u>	1070..2497 /gene="yigN" /note="synonym: b3832"
<u>CDS</u>	1070..2497 /gene="yigN" /function="phenotype; Not classified" /note="o475; 99 pct identical amino acid sequence and equal length to YIGN_ECOLI SW: P27850" /codon_start=1 /transl_table=11

```
/product="putative alpha helix chain"
/protein_id="AAC76835.1"
/db_xref="GI:1790266"
/translation="MDFSIMVYAVIALVGVAIGWLFASYQHAQQKAEQLAEREEMVAE
LSAAKQQITQSEHWRAECCELLNNEVRSLQSINTSLEADLREVTTTRMEAAQQHADDKIR
QMINSEQRLSEQFENLANRIFEHSNRRVDEQNRQSLNSLLSPLREQLDGFRQVQDSF
GKEAQERHTLTHEIRNLQQLNAQMAQEAINLTRALKGDNKTQGNWGEVVLTRVLEASG
LREGYEYETQVSIENDARSRMQPDVIVRLPQGDVVIDAKMTLVAYERFYNAEDDYTR
ESALQEHIASVRNHIRLLGRKDYQQLPGLRTL DYVLMFIPVEPAFLALDRQPELITE
ALKNNIMLVSPPTLLVALRTIANLWRYEHQSRNAQQIADRASKLYDKMRLFIDMSAI
GQSLDKAQDNYRQAMKKLSSGRGNVLAQAEAFRGLGVEIKREINPDLAEQAVSQDEEY
RLRSVPPEQPNDEAYQRDDEYNQQSR"

promoter 2451..2482
/gene="yigN"
/note="factor Sigma70; predicted +1 start at 4016339"

gene 2592..3347
/gene="ubiE"
/note="synonym: b3833"

CDS 2592..3347
/gene="ubiE"
/function="enzyme; Biosynthesis of cofactors, carriers:
Menaquinone, ubiquinone"
/note="o251; 100 pct identical to YIGO_ECOLI SW: P27851"
/codon_start=1
/transl_table=11
/product="2-octaprenyl-6-methoxy-1,4-benzoquinone -->
2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinone"
/protein_id="AAC76836.1"
/db_xref="GI:2367307"
/translation="MVDKSQETHFGFQTVAKEQKADMVAHVFSVASKYDVMNDLMS
FGIHRLWKRFTIDCSGVRRGQTVLDLAGGTGDLTAKFSRLVGETGKVVLADINESMPK
MGREKLNRNIGVIGNVEYVQANAEBALPFPDNTFDCITISFGLRNVTDKDKALRSMYRVL
KPGGRLLVLEFSKPIIEPLSKAYDAYSFHVLPRIGSLVANDADSYRYLAESIRMHDPQ
DTLKAMMQDAGFESVDYNNLTAGVVALHRGYKF"

gene 3361..3966
/gene="yigP"
/note="synonym: b3834"

CDS 3361..3966
/gene="yigP"
/function="orf; Unknown"
/note="o201; sequence changes cause internal changes to
thesequence relative to earlier versions (YIGP_ECOLI SW:
P27852) but the N- andC-terminal sequences are unaltered"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC76837.1"
/db_xref="GI:2367308"
/translation="MPFKPLVTAGIESLLNTFLYRSPALKTARSRLLGKVL RVEVKGF
STSLILVFSE RQVDVLGEWAGDADCTVIAYASVLPKLRDRQQLTALIRSGELEVVQGDI
QVVQNFVALADLAEFDP AELLAPYTG DIAAEGISKAMRGGAFLHHGIKQQRYVAEA
ITEEWRMAPGP LEVAWFAEETA AVERAVDALTKRLEKLEAK"

promoter 3858..3875
/gene="yigP"
/note="factor Sigma54; predicted +1 start at 4017732"

gene 3963..5603
/gene="yigR"
/note="synonym: b3835"

CDS 3963..5603
/gene="yigR"
```

/function="orf; Unknown"
/note="o546; sequence change joins ORFs yigQ, yigR, and yigS from earlier version (YIGQ_ECOLI SW: P27853; YIGR_ECOLI SW: P27854; YIGS_ECOLI SW: P27855); very good match to Azotobacter chroococcum partial ORF1 GB: U48404"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC76838.1"
/db_xref="GI:2367309"
/translation="MTPGEVRRLYFIIRTFLSYGLDELIPKMRITLPLRLWRYSLFWM
PNRHKDKLLGERLRLALQELGPVWIKFGQMLSTRDLFPPIADQLALLQDKVAPFDG
KLAKQQIEAAMGGLPVEAWFDDFEIKPLASASIAQVHTARLKSNGKEVVIKIRPDIL
PVIKADLKLIRLARWVPRLLPDGRRRLRPTEVVREYEKTLIDELNLLRESANAIQLRR
NFEDSPMLYIPEVYPDYCSEGMMVMERIYGIPVSDVAALEKNGTNMKLLAERGQVQFF
TQVFRDSFFHADMPGNIFVSYEHPENPKYIGIDCGIVGSLNKEDKRYLAENFIAFFN
RDYRKVAELHVDSGWVPPDTNVEEFFAIRTVCEPIFEKPLAEISFGHVLLNLFNTAR
RFNMEVQPQLVLLQKTLLEYVEGVGRQLYPQLDLWKTAKPFLESWIKDQVGIPALVRAF
KEKAPFWVEKMPPELPELVYDSLQCKYLOHSVDBKIARELQSNHVRQGSRYFLGIGAT
LVLSGTFLLVSRPEWGLMPGWL MAGGLIAWFGWRKTR"
gene 5640..5951
/gene="b3836"
CDS 5640..5951
/gene="b3836"
/function="orf; Unknown"
/note="o103; sequence change split ORF of earlier version; was part of yigT (o261; YIGT_ECOLI SW: P27856)"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC76839.1"
/db_xref="GI:2367310"
/translation="MRLCLIIYHRGTCMGGISIWQLLIIAIVVLLFGTKKLGSIGS
DLGASIKGFKKAMSDDEPKQDKTSQDADFTAKTIADKQADTNQEQAKTEDAKRHDKEQ
V"
gene 5815..6156
/gene="b3837"
CDS 5815..6156
/gene="b3837"
/function="putative factor; Not classified"
/note="o113; sequence change split ORF of earlier version; was part of yigT (o261; YIGT_ECOLI SW: P27856)"
/codon_start=1
/transl_table=11
/product="putative histone"
/protein_id="AAC76840.1"
/db_xref="GI:2367311"
/translation="MMNQSRIPVRMLILLRKLSPISRRIRIRNRLKQKTRSATIKSR
CNPCLISVLANCYWCSSASPPSGAATTACGGKNGSGLDSRVAFTGDNGAERTDPGVK
TPGVSGQSEKG"
gene 6035..6472
/gene="b3838"
CDS 6035..6472
/gene="b3838"
/function="orf; Unknown"
/note="o145; sequence change split ORF of earlier version; was part of yigT (o261; YIGT_ECOLI SW: P27856)"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"

gene /protein_id="AAC76841.1"
/db_xref="GI:2367312"
/translation="MAVKTVAGWIRALRSLATTVQNELTQELKLQEFQDSLKKVEKAS
LTNLTPELKASMDLRQAESMKRSYVANDPEKASDEAHTIHNFPVVKDNEAAHEGVTP
AAAQTQASSPEQKPETTPPEPVVKPAADAEPKTAAPSPSSSDKP"
6475..7251
/gene="yigU"
/note="synonym: b3839"
CDS 6475..7251
/gene="yigU"
/function="orf; Unknown"
/note="o258; sequence change joins two ORFs relative
toearlier version; 97.7 pct identical to the conceptual
ORF YIGU_ECOLI SW: P27857"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC76842.1"
/db_xref="GI:2367313"
/translation="MSVEDTQPLITHLIELRKRLNCIIAVIVIFLCLVYFANDIYHL
VSAPLIKQLPQGSTMATDVASPFPTPIKLTFMVSLILSAPVILYQVWAFIAPALYKH
ERRLVVPLLVSSSLFYIGMAFAYFVVFPLAFGLANTAPEGVQVSTDIASYSFVMA
LFMAFGVSFEVPVAIVLLCWMGITSPEDLRKKRPYVLVGAFVGMMLTTPDVFSQTL
AIPMYCLFEIGVFFSRFYVGKGRNREEENDAAEAESEKTEE"
gene 7281..7901
/gene="yigW_1"
/note="synonym: b3840"
CDS 7281..7901
/gene="yigW_1"
/function="orf; Unknown"
/note="o206; 98 pct identical to N-terminal 184 residues
of a 264 aa conceptual translation YIGW_ECOLI SW: P27859"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC76843.1"
/db_xref="GI:1790274"
/translation="MEYRMFDIGVNLTSQFAKDRDDVACAFDAGVNGLLITGTNLR
ESQQAQKLARQYSSCWSTAGVHPHDSSQWQAATEEAIIEELAAQPEVVAIGECGLDFNR
NFSTPEEQERAFVAQLRIAADLNMPVFMHCRDAHERFMTLLEPWLDKLPGAVLHCFTG
TREEMQACVAHGIYIGITGWVCDEPTRTGAAGTFAVDSGGKITDRN"
gene 7735..8076
/gene="yigW_2"
/note="synonym: b3841"
CDS 7735..8076
/gene="yigW_2"
/function="orf; Unknown"
/note="o113; 100 pct identical to 80 residues of a 264 aa
conceptual translation YIGW_ECOLI SW: P27859"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC76844.1"
/db_xref="GI:1790275"
/translation="MRFFIALPAHAKRCRRARMEFISALPVGFMNRRGLELRELLP
LIPAEKLLIETDAPYLLPRDLTPKPSSRRNEPAHLPHILQRIAHWRGEDAAWLAATTD
ANVKTLFGIAF"
gene complement(8073..8561)
/gene="rfaH"
/note="synonym: b3842"

CDS complement(8073..8561)
/gene="rfaH"
/function="regulator; Macromolecule metabolism:
Lipopolysaccharide"
/note="f162; 100 pct identical to RFAH_ECOLI SW: P26614;
alternate gene names sfrB, hlyT"
/codon_start=1
/transl_table=11
/product="transcriptional activator affecting biosynthesis
of lipopolysaccharide core, F pilin, and haemolysin"
/protein_id="AAC76845.1"
/db_xref="GI:1790276"
/translation="MQSWYLLYCKRGQLQRAQEHLERQAVNCLAPMITLEKIVRGKRT
AVSEPLFPNYLFVEFDPEVIHTTTINATRGVSHFVRFGASPAIVPSAVIHQLSVYKPK
DIVDPATPYPGDKVIITEGAFEGFQAIFTEPDGEARSMLLLNLINKEIKHSVKNTFR
KL"

promoter complement(8591..8619)
/note="factor Sigma70; predicted +1 start at 4022434"

promoter 8672..8704
/note="factor Sigma70; predicted +1 start at 4022561"

gene 8728..10221
/gene="yigC"
/note="synonym: b3843"

CDS 8728..10221
/gene="yigC"
/function="putative enzyme; Not classified"
/note="o497; ??? pct identical to conceptual
translationYIGC_ECOLI SW: P26615"
/codon_start=1
/transl_table=11
/product="putative oxidoreductase"
/protein_id="AAC76846.1"
/db_xref="GI:1790277"
/translation="MDAMKYNDLRDFLTLLLEQQGELKRITLPVDPHLEITEIADRTL
RAGGPALLFENPKGYSMPLVLCNLFGTPKRVAMGMGQEDVSALREVGKLLAFLKEPEPPK
GFRDLFDKLPQFKQVLNMPKRLRGAPCQQKIVSGDDVDLNRIPIMTCWPEDAAPLIT
WGLTVTRGPHKERQNLGIYRQQLIGKNKLIMRWLSHRGGALDYQEWCAAHPGERFPVS
VALGADPATILGAVTPVPDTLSEYAFAGLLRGTKTEVVKICISNDLEVPASAEIVLEGY
IEQGETAPEGPYGDHTGYNEVDSFPVFTVTHITQREDAIYHSTYTGRPPDEPAVLGV
ALNEVFVPILOKQFPEIVDFYLPPEGCSYRLAVVTIKKQYAGHAKRVMGMVWSFLRQF
MYTKFVIVCDDVDNARDWNDVIWAITTRMDPARDTVLVENTPIDYLDFAFPVSGLSK
MGLDATNKWPGETQREWGRPIKDPDVVAHIDAIWDELAIFNNGKSA"

gene 10267..10968
/gene="ubiB"
/note="synonym: b3844"

CDS 10267..10968
/gene="ubiB"
/EC_number="1.6.8.-"
/function="enzyme; Energy metabolism, carbon: Electron
transport"
/note="o233; 99 pct identical amino acid sequence and
equal length to UBIB_ECOLI SW: P23486; similar to
monooxygenase subunits fromMethylococcus capsulatus and
Pseudomonas putida"
/codon_start=1
/transl_table=11
/product="ferrisiderophore reductase; flavin reductase
(NADPH:flavin oxidoreductase)"
/protein_id="AAC76847.1"
/db_xref="GI:2367314"

/translation="MTTLSCKVTSVEAITDTVYRVRIVPDAAFSFRAGQYLMVVMDER
DKRPFSMASTPDEKGFIELHIGASEINLYAKAVMDRILKDHQIVVDIPHGEAWLRDDE
ERPMLIAGGTGFSYARSILLTALARNPNRDITYWGGREEQHLYDLCELEALSLKHP
GLQVVPVVEQPEAGWRGRTGTVLTAVLQDHGTLAEHDIYIAGRFE MAKIARDLFCSE
NAREDR LFGDAFAFI"

BASE COUNT 2621 a 2682 c 3073 g 2681 t

ORIGIN

```
1 agtcctgcag aatgaagggt gatttatgtg atttgcac ca tttttgggtg gtaaatttat
61 gcaacgcatt tgcgtcatgg tgatgagtat cacgaaaaaa tgttaaacc ttcggtaaag
121 tgtctttttg cttcttctga ctaaaccgat tcacagagga gttgtatatg tccaagtctg
181 atgtttttca tctcggcctc actaaaaacg atttacaagg ggctacgctt gccatcgctc
241 ctggcgaccc ggatcgtgtg gaaaagatcg ccgcgctgat ggataagccg gttaaagctg
301 catctcaccg cgaattcact acctggcgtg cagagctgga tggtaaacct gttatcgctc
361 gctctaccgg tatcggcggc ccgtctacct ctattgctgt tgaagagctg gcacagctgg
421 gcattcgcac cttcctgcgt atcgggtacaa cgggcgctat tcagccgcat attaattgtg
481 gtgatgtcct ggttaccacg gcgtctgtcc gtctggatgg cgcgagcctg cacttcgcac
541 cgctggaatt cccggctgtc gctgatttcg aatgtacgac tgcgctgggt gaagctgcga
601 aatccattgg cgcgacaact cacgttggcg tgacagcttc ttctgatacc ttctaccag
661 gtcaggaacg ttacgatact tactctggtc gcgtagttcg tcactttaaa ggttctatgg
721 aagagtggca ggcgatgggc gtaatgaact atgaaatgga atctgcaacc ctgctgacca
781 tgtgtgcaag tcagggcctg cgtgccggtg tggtagcggg tgttatcggt aaccgcaccc
841 agcaagagat cccgaatgct gagacgatga aacaaaccga aagccatgcg gtgaaaatcg
901 tgggtggaagc ggcgcgctcg ctgctgtaat tctcttctcc tgtctgaagg ccgacgcgtt
961 cggccttttg tatttttgcg tagcgctcg caggaaatgc ctttccaact ggacgtttgt
1021 acagcacaat tctattttgt gcgggtaagt tgttgcgtca ggaggcgttg tggatttctc
1081 aatcatgggt tacgcagtta ttgcgttggt ggggtgaggc attggctggc tgtttgccag
1141 ttatcaacat gcgcagcaaa aagccgagga attagctgaa cgtgaagaga tggtcgcgga
1201 gttaagcgcg gcaaaacaac aaattacca aagcgagcac tggcggtgcg agtgcgaggt
1261 actcaataac gaagtgcgca gcctgcaaag tattaacacc tctctggagg ccgatctgcg
1321 tgaagtaacc acgcggatgg aagccgcaca gcaacatgct gacgataaaa ttcgccagat
1381 gattaacagc gagcagcgcc tcagtgcgca gtttgaaaac ctgcaccaac gtatttttga
1441 gcacagcaat cgccgggttg atgagcaaaa ccgtcagagt ctgaacagcc tgttgcgcc
1501 gctacgtgaa caactggacg gtttcgcgcg tcaggttcag gacagcttcg gtaaagaagc
1561 acaagaacgc cataccctga cccacgaaat tcgcaatctc cagcaactca acgcgcaaat
1621 ggcccaggaa gcgatcaacc tgacgcgcgc gctgaaaggc gacaataaaa cccagggcaa
1681 ctggggcgag gtagtattga cgcgggtgct ggaggcttcc ggtctgcgtg aaggggtatg
1741 atatgaaacc caggtcagca tcgaaaatga cgcgcgctcg cggatgcagc cggatgtcat
1801 cgtgcgcctg ccgcagggaa aagatgtggg gatcgacgcc aaaatgacgc tggtcgccta
1861 tgaacgctat tttaacgccg aagacgacta caccgcgaa agcgcgtac aggaacatat
1921 cgcgctcggg cgtaaccata tccgtttgct gggacgcaaa gattatcaac agctgccggg
1981 gctgcgaact ctggattacg tgcgtgattt tattcccggt gaaccgcgtt ttttactggc
2041 gcttgaccgc cagccggagc tgatcccgca agcgttgaaa aacaacatca tgctgggtag
2101 cccgactacg ctgctggtgg cgctgcgcac tatcgccaac ctgtggcggt atgagcatca
2161 aagccgcaac gccagcaaaa tcgcccagcg tgccagcaag ctgtacgaca agatgcgttt
2221 gttcatcgat gacatgtccg cgattggtca aagtctcgac aaagcgcagg ataattatcg
2281 gcaggcaatg aaaaaactct cttcaggggc cggaaatgtg ctggcgcagg cagaagcggt
2341 tcgcggttta ggagtagaaa ttaaaccgca gattaatccg gatttggctg aacaggcggt
2401 gagccaggat gaagagtatc gacttcggtc ggttcgggag cagccgaatg atgaagctta
2461 tcaacgcgat gatgaatata atcagcagtc gcgctagccc attgggagta gttaagccgg
2521 gtagaaatct agggcatcga cgcaccaatc gttacacttc tggaaacaatt ttttgatgag
2581 caggcattga gatggtggat aagtcacaag aaacgacgca ctttggtttt cagaccgtcg
2641 cgaaggaaca aaaagcggat atggtcgccc acgttttcca ttccgtggca tcaaaatacg
2701 atgtcatgaa tgatttgatg tcatttggtg ttcacgtttt gtggaagcga ttcacgattg
2761 attgcagcgg cgtacgccgt gggcagaccg tgctggatct ggctggggg accggcgacc
2821 tgacagcgaa attctccgcg ctggtcggag aaactggcaa agtggctcctt gctgatatca
2881 atgaatccat gcccaaaatg ggccgcgaga agctgcgtaa tatcggtgtg attggcaacg
2941 ttgagtatgt tcaggcgaac gctgaggcgc tgccgttccc ggataacacc tttgattgca
3001 tcaccatttc gtttgggtct cgtaacgtca ccgacaaaga taaagcactg cgttcaatgt
3061 atcgcgctgt gaaaccgggc ggccgcctgc tgggtgcttg gttctcgaag ccaattatcg
3121 agccgctgag caaagcctat gatgcatact ctttccatgt gctgcgcgct attggctcac
```

```
3181 tggctcgcgaa cgacgccgac agctaccggt atctggcaga atccatccgt atgcatcccc
3241 atcaggatac cctgaaagcc atgatgcagg atgccggatt cgaaagtgtc gactactaca
3301 atctgacggc aggggttgtg gcgctgcac gtggttataa gttctgacag gagaccggaa
3361 atgcctttta aacctttagt gacggcagga attgaaagtc tgctcaacac cttcctgtat
3421 cgctcaccocg cgctgaaaac ggcccgcctc cgtctgctgg gtaaaagtatt gcgctggag
3481 gtaaaaggct tttcgacgtc attgattctg gtgttcagcg aacgccaggt tgatgtactg
3541 ggcgaaatggg caggcgatgc tgactgcacc gttatcgctt acgccaggtg gttgccgaaa
3601 cttcgcgatc gccagcagct taccgcactg attcgcagtg gtgagctgga agtgaggggc
3661 gatattcagg tgggtgcaaaa cttcgtttcg ctggcagatc tggcagagtt cgaccctgcg
3721 gaactgctgg ccccttatac cgggtgatat gccgctgaag gaatcagcaa agccatgcgc
3781 ggaggcgcaa agttcctgca tcacggcatt aagcgccagc aacgttatgt ggcggaagcc
3841 attactgaag agtggcgtat ggcacccggt ccgcttgaag tggcctggtt tgcggaagag
3901 acggctgccg tcgagcgtgc tgttgatgcc ctgaccaaac ggctggaaaa actggaggct
3961 aatgacgcc aggtgaagta cggcgcttat atttcatcat tcgcactttt ttaagctacg
4021 gacttgatga actgatcccc aaaatgcgta tcaccctgcc gctacggcta tggcgatact
4081 cattattctg gatgccaat cggcataaag acaaaactttt aggtgagcga ctacgactgg
4141 ccctgcaaga actggggccg gtttggtatca agttcgggca aatgttatca acccgccgcg
4201 atctttttcc accgcatatt gccgatcagc tggcgttatt gcaggacaaa gttgctccgt
4261 ttgatggcaa gctggcgaa gacagattg aagctgcaat gggcggttg ccggtagaag
4321 cgtggtttga cgattttgaa atcaagccgc tggcttctgc ttctatcgcc caggttcata
4381 ccgcgcgatt gaaatcgaat ggtaaagagg tggtgattaa agtcatccgc ccgcatattt
4441 tgccggttat taaagcggat ctgaaactta tctaccgtct ggctcgctgg gtgccgcgtt
4501 tgctgccgga tggctgccgt ctgcccga ccgaagtggg gcgcgagtac gaaaagacat
4561 tgattgatga actgaatttg ctgcccgaat ctgccaacgc cattcagctt cggcgcaatt
4621 ttgaagacag ccgatgctc tacatcccgg aagtttacc tgactattgt agtgaaggga
4681 ttgatggtgat ggagcgcat tacggcattc cgggtgctga tgttgccggt ggggagaaaa
4741 acggcactaa catgaaattg ctggcggaac gcggcgtgca ggtgttctt actcaggtct
4801 ttcgcgacag ctttttccat gccgatatgc accctggcaa catcttcgta agctatgaac
4861 acccggaaaa ccggaatat atcggcattg attgcgggat tgttggtcgt ctaaacaaaag
4921 aagataaacg ctatctggca gaaaacttta tcgccttctt taatcgcgac tatcgcaaag
4981 tggcagagct acacgtcgat tctggctggg tgccaccaga taccaacgtt gaagagttcg
5041 aatttgccat tcgtacggtc tgtgaacctt tctttgagaa accgctggcc gaaatttcgt
5101 ttggacatgt actgttaaat ctgtttaata cggcgctcgc cttcaatatg gaagtgcagc
5161 cgcaactggg gttactccag aaaaccctgc tctacgtcga aggggtagga cgccagcttt
5221 atccgcaact cgatttatgg aaaacggcga agcctttcct ggagtcgtgg attaaagatc
5281 aggtcgggtat tcttcgctg gtgagagcat ttaaagaaaa agcgccgttc tgggtcgaaa
5341 aaatgccaga actgcctgaa ttggtttacg acagtttgcg ccagggcaag tatttacagc
5401 acagtgttga taagattgcc cgcgagcttc agtcaaatca tgtacgtcag ggacaatcgc
5461 gttattttct cggaattggc gctacgttag tattaagtgg cacattcttg ttggtcagcc
5521 gacctgaatg ggggctgatg cccggctggg taatggcagg tggctctgat gcctggtttg
5581 tcggttggcg caaaacacgc tgattttttc atcgctcaag gcgggcccgt taacgtataa
5641 tgcggctttg tttaatcatc atctaccaca gaggaacatg tatgggtggg atcagtattt
5701 ggcagttatt gattattgcc gtcactgttg tactgctttt tggcaccaaa aagctcggct
5761 ccactcgggt cgactttggg gcgtcgatca aaggctttaa aaaagcaatg agcgatgatg
5821 aaccaaagca ggataaaacc agtcaggatg ctgattttac tgcgaaaaact atcgccgata
5881 agcaggcgga tacgaatcag gaacaggcta aaacagaaga cgcgaagcgc cagcataaag
5941 agcaggtgta atccgtgttt gatatcggtt ttagcgaact gctattggtg ttcacatcgc
6001 gcctcgccgc cttctggggc cgcaacgact gcctgtggcg gtaaaaaacg tagcgggctg
6061 gattcgcgcg ttgcgttcac tggcgacaac ggtgcagaac gaactgaccc aggagttaaa
6121 actccaggag tttcaggaca gtctgaaaaa ggttgaaaag gcgagcctca ctaacctgac
6181 gcccgaactg aaagcgtcga tggatgaact acgccaggcc gcggagtcga tgaagcgttc
6241 ctacgttgca aacgatcctg aaaaggcgag cgatgaagcg cacaccatcc ataaccgggt
6301 ggtgaaagat aatgaagctg cgcgatgggg cgtaacgcct gccgctgcac aaacgcaggc
6361 cagttcgccg gaacagaagc cagaaaccac gccagagccg gtggtaaaaa ctgctgcgga
6421 cgctgaaccg aaaaccgctg caccttcccc ttcgtcgagt gataaaaccgt aaacatgtct
6481 gtagaagata ctcaaccgct tatcacgcat ctgattgagc tgcgtaagcg tctgctgaac
6541 tgcattatcg cggtgatcgt gatattcctg tgtctggtct atttcgcaa tgacatctat
6601 cacctgggtat ccgcgccatt gatcaagcag ttgccgcaag gttcaacgat gatcgccacc
6661 gacgtggcct cgccgttctt tacgccgatc aagctgacct ttatggtgtc gctgattctg
6721 tcagcgccgg tgattctcta tcaggtgtgg gcatttatcg ccccgagcgt gtataagcat
```

```
6781 gaacgtcgcc tgggtggcgcc gctgctgggt tccagctctc tgctgtttta taccggcatg
6841 gcattcgccct accttgtgggt ctttccgctg gcatttgggt tccttgccaa taccgcgcgcg
6901 gaaggggtgac aggtatccac cgacatcgcc agctatttaa gcttcgttat ggcgctgttt
6961 atggcgtttg gtgtctcctt tgaagtgcgc gtagcaattg tgctgctgtg ctggatgggg
7021 attacctcgc cagaagactt acgcaaaaaa cggcgtatg tgctggttgg tgcattcggt
7081 gtcgggatgt tgctgacgcc gccggatgtc ttctcgcaaa cgctgttggc gatcccgatg
7141 tactgtctgt ttgaaatcgg tgtcttcttc tcacgctttt acgttggtaa agggcgaaat
7201 cgggaagagg aaaacgacgc tgaagcagaa agcgaaaaaa ctgaagaata aattcaaccg
7261 cccgtcaggg cggttgtcat atggagtaca ggatgtttga tatcggcggt aatttgacca
7321 gttcgcgaatt tgcgaagac cgtgatgatg ttgtagcgtg cgctttttgac gcgggagtta
7381 atgggctact catcaccggc actaacctgc gtgaaagcca gcaggcgcaa aagctggcgc
7441 gtcagtattc gtcctgttgg tcaacggcgg cggtacatcc tcacgacagc agccagtggc
7501 aagctgcgac tgaagaagcg attattgagc tggccgcgca gccagaagtg gtggcgattg
7561 gtgaatgtgg tctcgacttt aaccgcaact ttctgacgcc ggaagagcag gaacgcgctt
7621 ttgttgccca gctacgcatt gccgcagatt taaacatgcc ggtatttatg cactgtcgcg
7681 atgccacga gcggtttatg acattgctgg agccgtgggt ggataaactg cctggtgcgg
7741 ttcttcattg ctttaccggc acacgcgaag agatgcaggc gtgctggcg catggaattt
7801 atatcgccat taccggttgg gtttgcgatg aaccgacgcg gactggagct gcgggaactt
7861 ttgcccgtga ttccggcgga aaaattactg atcgaaactg atgcgccgta tctgctccct
7921 cgcgatctca cgccaaagcc atcatcccgg cgcaacgagc cagcccatct gccccatatt
7981 ttgcaacgta ttgcgcaact gcgtggagaa gatgccgcat ggctggctgc caccacggat
8041 gctaattgtca aaacactggt tgggatttgc ttttagagtt tgcggaactc ggtattcttc
8101 aactgtgct taatctcttt attaataaga ttaagcaata gcatggagcg agcctcacca
8161 tcgggttcgg tgaaaatggc ctgaaagcct tcgaacgcgc cttcggtaat aatcacctta
8221 tcacccggat aaggggttgc cggatcgaca atgtctttcg gtttatatac cgatagctga
8281 tgaataaccg ccgatgggac tatcgctggc gacgcgcaa agcgcacgaa gtggctgaca
8341 ccgcggtcgc cgttgatagt cgtggtatga atcacttctg ggtcaaattc cacaacacgg
8401 tagttggggc acaatggctc actgactgca gtacgttttc cagcacgat ttttccagg
8461 gtgatcatcg gtgccaggca attcacagcc tgtctttcga ggtgttcctg ggcacgttga
8521 agttgcccgc gcttgcacta cagtaaatac caggattgca taatgactct tatccgttta
8581 atcggggcgc aaggatagca aaagctttac gctaaagtta ttatattccc cggtttgcgt
8641 tataccgtca gagttcacgc taatttaaca aatttacagc atcgcaaaga tgaacgccgt
8701 ataatgggcg cagattaaga ggctacaatg gacgccatga aatataacga ttacgcgac
8761 ttcttgacgc tgcttgaaca gcagggtgag ctaaaacgta tcacgctccc ggtggatccg
8821 catctggaat tcaactgaaat tgctgaccgc actttgcgtg ccggtggggc tgcgctgttg
8881 ttcgaaaacc ctaaaggcta ctcaatgccg gtgctgtgca acctgttcgg tacgcaaag
8941 cgcgtggcga tgggcatggg gcaggaagat gtttcggcgc tgcgtgaagt tggtaaatta
9001 ttggcgtttc tgaaagagcc ggaagccgca aaagggttcc gcgacctgtt tgataaactg
9061 ccgcagttta agcaagtatt gaacatgccg acaaagcggc tgcgtggtgc gccctgccaa
9121 caaaaaatcg tctctggcga tgacgtcgat ctcaatcgca ttcccattat gacctgctgg
9181 ccggaagatg ccgcgccgct gattacctgg gggctgacag tgacgcgcgg ccacataaa
9241 gagcggcaga atctgggcat ttatcgccag cagctgattg gtaaaaacaa actgattatg
9301 cgctggctgt cgcacgcggc cggcgcgctg gattatcagg agtgggtgtc ggcgcacccg
9361 ggcgaacgtt tcccggtttc tgtggcgctg ggtgcccgat ccgccacgat tctcggtgca
9421 gtcactcccc ttccggatac gctttcagag tatgcgtttg ccggtattgc acgtggcacc
9481 aagaccgaag tggatgaagt tatctccaat gatcttgaag tgcccgcag tgcggagatt
9541 gtgctggaag ggtatatcga acaaggcgaa actgcgcggg aaggggcgta tggcgaccac
9601 accggttact ataatgaagt cgatagtttc ccggtattta ccgtgacgca tattaccag
9661 cgtgaagatg cgatttacca ttccacctat accgggcgtc cgccagatga gccgcgggtg
9721 ctgggtgtcg cactgaacga agtgtttgtg ccgattctgc aaaaacagtt ccggaaatt
9781 gtcgattttt acctgccgcc ggaaggctgc tcttatcgcc tggcggtagt gacaatcaaa
9841 aacagtagc ccggacacgc gaagcgcgtc atgatggcg tctggctggt cttacgccag
9901 tttatgtaca ctaaatgtgt gatcgtttgc gatgatgacg ttaacgcacg cgactggaac
9961 gatgtgattt gggcgattac caccgctatg gaccggcgcc gggatactgt tctggtagaa
10021 aatacgcta ttgattatct ggattttgcc tcgcctgtct ccgggctggg ttcaaaaatg
10081 gggctggatg ccacgaataa atggccgggg gaaaccacgc gtgaatgggg acgtcccatc
10141 aaaaagatc cagatgttgt cgcgcataat gacgccatct gggatgaact ggctattttt
10201 aacaacggta aaagcgccct atgcgcgttt gttttgccct atttatcgat ccgacagaga
10261 aagcgcagta caaccttaag ctgtaaagtg acctcggtag aagctatcac ggataccgta
10321 tatcgtgtcc gcacgtgcc agacgcggcc ttttcttttc gtgctggtca gtatttgatg
```

```
10381 gtagtgatgg atgagcgcgga caaacgtccg ttctcaatgg cttcgacgcc ggatgaaaaa
10441 ggggtttatcg agctgcatat tggcgcttct gaaatcaacc tttacgcgaa agcagtcatg
10501 gaccgcatcc tcaaagatca tcaaactcgtg gtcgacattc cccacggaga agcgtggctg
10561 cgcgatgatg aagagcgctc gatgattttg attgcgggcg gcaccgggtt ctcttatgcc
10621 cgctcgattt tgctgacagc gttggcgcggt aacccaaacc gtgatatac catttactgg
10681 ggcgggctg aagagcagca tctgtatgat ctctgcgagc ttgaggcgct ttcgttgaag
10741 catcctggtc tgcaagtggg gccggtggtt gaacaaccgg aagcgggctg gcgtgggctg
10801 actggcaccg tgtaacggc ggtattgcag gatcacggtg cgctggcaga gcatgatatac
10861 tatattgccg gacgttttga gatggcgaaa attgcccgcg atctgttttg cagtgagcgt
10921 aatgcgcggg aagatcgctt gtttggcgat gcgtttgcat ttatctgaga tataaaaaaa
10981 cccgcccctg acaggcggga agaacggcaa ctaaactgtt attcagtggc atttagatct
11041 atgacgtatc tggcaaa
```

//

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

May 20 2003 11:20:12



GGTCAGGATACGACTTCGACCGCTAGAGGATCGGATCCCGGCGATTATATAGCTCGATCGTC
 TTCTCATATACCGCGATATGGGATATATACACACACGCGGCGGATAGCATGACTGATCT
 CCGCATCTACCGCTACCGCTACCGCTACCGCTACCGCTACCGCTACCGCTACCGCTACCGCT
 CACAGACTACCGCTACCGCTACCGCTACCGCTACCGCTACCGCTACCGCTACCGCTACCGCT

PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

OMIM

Boo

Search for

Limits

Preview/Index

History

Clipboard

Details

Show:

1: AE000167. Escherichia coli ...[gi:1786836]

Links

LOCUS AE000167 10264 bp DNA linear BCT 01-DEC-2000
 DEFINITION Escherichia coli K12 MG1655 section 57 of 400 of the complete genome.
 ACCESSION AE000167 U00096
 VERSION AE000167.1 GI:1786836
 KEYWORDS
 SOURCE Escherichia coli K12
 ORGANISM Escherichia coli K12
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
 REFERENCE 1 (bases 1 to 10264)
 AUTHORS Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y.
 TITLE The complete genome sequence of Escherichia coli K-12
 JOURNAL Science 277 (5331), 1453-1474 (1997)
 MEDLINE 97426617
 PUBMED 9278503
 REFERENCE 2 (bases 1 to 10264)
 AUTHORS Blattner, F.R.
 TITLE Direct Submission
 JOURNAL Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
 Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459
 REFERENCE 3 (bases 1 to 10264)
 AUTHORS Blattner, F.R.
 TITLE Direct Submission
 JOURNAL Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
 Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459
 REFERENCE 4 (bases 1 to 10264)
 AUTHORS Plunkett, G. III.
 TITLE Direct Submission
 JOURNAL Submitted (13-OCT-1998) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
 COMMENT This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants HG00301 and HG01428 (from the Human Genome Project and NCHGR). The entire sequence was independently determined from E. coli K12 strain MG1655. Predicted open reading frames were determined using GeneMark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 [e-mail: mark@amber.gatech.edu]. Open reading frames that have been correlated with genetic loci are being annotated with CG

Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (<http://cgsc.biology.yale.edu>). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (<http://www.genetics.wisc.edu>). *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.

FEATURES	Location/Qualifiers
<u>source</u>	1..10264 /organism="Escherichia coli K12" /mol_type="genomic DNA" /strain="K12" /sub_strain="MG1655" /db_xref="taxon:83333"
<u>promoter</u>	64..92 /note="factor Sigma70; predicted +1 start at 651381"
<u>gene</u>	176..1834 /gene="citA" /note="synonym: b0619"
<u>CDS</u>	176..1834 /gene="citA" /function="putative regulator; Not classified" /note="o552 was o455; This 455 aa ORF is 29 pct identical (15 gaps) to 441 residues of an approx. 544 aa protein YJDH_ECOLI SW: P39272; This 77 aa ORF is 28 pct identical (2 gaps) to 59 residues of an approx. 400 aa protein YWFF_BACSU SW: P39642" /codon_start=1 /transl_table=11 /product="putative sensor-type protein" /protein_id="AAC73720.1" /db_xref="GI:1786837" /translation="MLQLNENKQFAFFQRLAFPLRIFLLILVFSIFVIAALAQYFTAS FEDYLTLHVRDMAMNQAKIIASNDSVISAVKTRDYKRLATIANKLQRTDFDYVVIGD RHSIRLYHPNPEKIGYPMQFTKQGALEKGESYFITGKSGMGMAMRAKTPIFDDDGKVI GVVSIGYLVSKIDSWRAEFLLPMAGVFVLLGILMLLSWFLAAHIRRQMMGMEPKQIA RVVRQQEALFSSVYEGLIAVDPHGYITAINRNARKMLGLSSPGRQWLGPIVEVVRPA DFFTEQIDEKRQDVVANFNGLSVIANREAIRSGDDLGAIIISFRSKDEISTLNAQLTQ IKQYVESLRTLRLHEHLNWMSTLNGLLQMKEYDRVLAMVQGESQAQQQLIDSLREAFAD RQVAGLLFGKVQRARELGLKMIIVPGSQLSQLPPGLDSTEFAAIVGNLLDNAFEASLR SDEGNKIVELFLSDEGDDVVIEVADQGCVPESLRDKIFEQGVSTRADEPGEHGIGLY LIASYVTRCGGVITLEDNDPCGTLFSIYIPKVKPNDSSINPIDR"
<u>gene</u>	1803..2483 /gene="citB" /note="synonym: b0620"
<u>CDS</u>	1803..2483 /gene="citB" /function="putative regulator; Not classified" /note="o226; 100 pct identical to Shigella flexneri"

protein GB: SFU29654_1 ACCESSION: U29654"
/codon_start=1
/transl_table=11
/product="sequence similarity to Shigella regulator"
/protein_id="AAC73721.1"
/db_xref="GI:1786838"
/translation="MTAPLTLLIVEDETPLAEMHAEYIRHIPGFSQILLAGNLAQARM
MIERFKPGLILLDNYLPDGRGINLLHELVAHYPGDVVFTTAASDMETVSEAVRCGVF
DYLIKPIAYERLGQTLTRFRQRKHMLESIDSASQKQIDEMFNAYARGEPAKDELPTGID
PLTLNAVRKLFKEPGVQHTAETVAQALTISRRTARRYLEYCASRHLIAEIVHGKVGR
PQRIYHSG"
gene complement(2524..3909)
/gene="dcuC"
/note="synonym: b0621"
CDS complement(2524..3909)
/gene="dcuC"
/function="transport; Transport of small molecules:
Carbohydrates, organic acids, alcohols"
/note="f461; This 461 aa ORF is 36 pct identical (6 gaps)
to 441 residues of an approx. 456 aa protein YHCL_ECOLI
SW: P45428"
/codon_start=1
/transl_table=11
/product="transport of dicarboxylates"
/protein_id="AAC73722.1"
/db_xref="GI:1786839"
/translation="MLTFIELLLIGVVVIVGVARYIIKGYSATGVLFVGGLLLLLIISAI
MGHKVLPSSQASTGYSATDIVEYVKILLMSRGGDLGMMIMMLCGFAAYMTHIGANDMV
VKLASKPLQYINSPYLLMIAAYFVACLSLAVSSATGLGVLLMATLFPVMVNVGISRG
AAAAICASPAAILAPTSGDVVLAQAQASEMSLIDFAFKTTLPISIAAIGMAIAHFFW
QRYLDKKEHISHEMLDVSEITTTAPAFYAILPFTPIIGVLI FDGKWGPQLHIITILVI
CMLIASILEFLRSFNTQKVFSGLEVAYRGMADAFANVVMMLVAAGVFAQGLSTIGFIQ
SLISIATSFSGSASIILMLVLVILTMLAAVTTGSGNAPFYAFVEMIPKLAHSSGINPAY
LTIPMLQASNLGRTLSPVSGVVVAVAGMAKISPFEVVKRTSVPVLVGLVIVIVATELM
VPGTAAAVTGK"
promoter complement(3939..3967)
/note="factor Sigma70; predicted +1 start at 655214"
protein_bind complement(4051..4076)
/note="central position to predicted promoter: -107"
/bound_moiety="ArsR predicted site"
promoter 4344..4372
/note="factor Sigma70; predicted +1 start at 655661"
promoter 4443..4471
/note="factor Sigma70; predicted +1 start at 655760"
gene 4498..5058
/gene="crcA"
/note="synonym: b0622"
CDS 4498..5058
/gene="crcA"
/function="orf; Unknown"
/note="o186; 100 pct identical to GB: ECOGICA_1 ACCESSION:
L29054; 100 pct. identical to 156 amino acids of
YBEG_ECOLI SW: P37001 but contains 30 additional
N-terminal residues"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC73723.1"
/db_xref="GI:1786840"
/translation="MNVSKYVAIFS FVFIQLISVGKVFANADEWMTTFRENIAQTWQQ

PEHYDLYIPAITWHARFAYDKEKTDYNERPWGGGFGLSRWDEKGNWHGLYAMAFKDS
WNKWEPIAGYGWESTWRPLADENFHLGLGFTAGVTARDNWNYPVLLPLASVGYGP
VTFQMTYIPGTYNNGNVYFAWMRFQF"

promoter 5084..5112
/note="factor Sigma70; predicted +1 start at 656401"

gene 5233..5442
/gene="cspE"
/note="synonym: b0623"

CDS 5233..5442
/gene="cspE"
/function="phenotype; Not classified"
/note="o69; 100 pct identical to CSPE_ECOLI SW: P36997"
/codon_start=1
/transl_table=11
/product="cold shock protein"
/protein_id="AAC73724.1"
/db_xref="GI:1786841"
/translation="MSKIKGNVWKFNESKGFGFITPEDGSKDVFVHFSAIQTNGFKTL
AEGQRVFEFEITNGAKGPSAANVIAL"

gene complement(5496..5879)
/gene="crcB"
/note="synonym: b0624"

CDS complement(5496..5879)
/gene="crcB"
/function="orf; Unknown"
/note="f127; formerly designated ybeI"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC73725.1"
/db_xref="GI:1786842"
/translation="MLQLLLAVFIGGGTGSVARWLLSMRFNPLHQAIPLGTLTANLIG
AFIIGIGFAWFSRMTNIDPVWVKVLITTGFCGGLTTFSTFSAEVVFLQEGRFGWALLN
VFNLLGSFAMTALAFWLFSASTAH"

promoter 5870..5899
/note="factor Sigma70; predicted +1 start at 657188"

promoter complement(5945..5972)
/note="factor Sigma70; predicted +1 start at 657220"

gene 5972..6199
/gene="ybeH"
/note="synonym: b0625"

CDS 5972..6199
/gene="ybeH"
/function="orf; Unknown"
/note="o75; 100 pct identical to YBEH_ECOLI SW: P39874"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC73726.1"
/db_xref="GI:1786843"
/translation="MLVAAGQFAVTSVWEKNAEICASLMAQAAENDASLFALPEALLA
RDDHDADLSVKSAQLLEGEFLGLYGKVNVT"

gene 6196..6759
/gene="ybeM"
/note="synonym: b0626"

CDS 6196..6759
/gene="ybeM"
/function="putative enzyme; Not classified"
/note="o187; This 187 aa ORF is 40 pct identical (33 gaps)
to 165 residues of an approx. 312 aa protein YJM6_YEAST"

SW: P47016"
/codon_start=1
/transl_table=11
/product="putative amidase"
/protein_id="AAC73727.1"
/db_xref="GI:1786844"
/translation="MMTTILTIHVPSTPGRAWNMLVALQAGNIVARYAKLHLYDAFAI
QESRRVDAGNEIAPLLEVEGMKVGMLTCYDLRFPELALAQALQGAEILVLPAAWVRGP
LKEHHWSTLLAARALDTTTCYMAAGECGNKNIGQSRIIDPFGVTIAAASEMPALIMAE
VTPERVQRVRAQLPVLNNRRFAPPQLL"

promoter 6777..6806
/note="factor Sigma70; predicted +1 start at 658095"

gene 6888..7091
/gene="ybeC"
/note="synonym: b0627"

CDS 6888..7091
/gene="ybeC"
/function="orf; Unknown"
/note="ø67; 98 pct identical to YBEC_ECOLI SW: P25895"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC73728.1"
/db_xref="GI:1786845"
/translation="MGEISITKLLVVAALVVLLFGTKKLRTLGGDLGAAIKGFKKAMN
DDDAAAKKGADVLDQAEKLSHKE"

gene complement(7192..8157)
/gene="lipA"
/note="synonym: b0628"

CDS complement(7192..8157)
/gene="lipA"
/function="putative enzyme; Biosynthesis of cofactors,
carriers: Lipoate"
/note="f321; 99 pct identical to LIPA_ECOLI SW: P25845"
/codon_start=1
/transl_table=11
/product="lipoate synthesis, sulfur insertion?"
/protein_id="AAC73729.1"
/db_xref="GI:1786846"
/translation="MSKPIVMERGVKYRDADKMALIPVKNVATEREALLRKPEWMKIK
LPADSTRIQGIKAAMRKNGLHSVCEEASCPNLAECFNHGTATFMILGAICTRRCPFCD
VAHGRPVAPDANEPVKLAQTIADMALRYVVITSVDRDLDLDGGAQHFAADCITAIREKS
PQIKIETLVPDFRGRMDRALDILTATPPDVFNHNLENVPRIYRQVRPGADYNWSLKL
ERFKEAHPEIPTKSGLMVGLGETNEEIIIEVMRDLRRHGVMTLTLGQYLQPSRHHLVPQ
RYVSPDEFDEMKAELAMGFTHAACGPFVRSSYHADLQAKGMEVK"

promoter complement(8251..8280)
/note="factor Sigma70; predicted +1 start at 659526"

gene complement(8366..9166)
/gene="ybeF"
/note="synonym: b0629"

CDS complement(8366..9166)
/gene="ybeF"
/function="putative regulator; Not classified"
/note="f266; This 266 aa ORF is 99 pct identical (1 gap)
to 266 residues of an approx. 320 aa protein YBEF_ECOLI
SW: P30979"
/codon_start=1
/transl_table=11
/product="putative transcriptional regulator LYSR-type"
/protein_id="AAC73730.1"

gene /db_xref="GI:1786847"
/translation="MLNLTPSAISQSIQKLRVIFPDPLFIRKQGQVTPAFAMHLHEY
ISQGLSILGALDIEGSYDKQRTITITATTPSVGALVLPVIYRAIKTHYPQLLLRNPP
SDAENQLSQFQTDLIIDNMFCTNRTVQHHVLFDTNMVLCIREGNPLLSLEDDRETIDN
AAHVLLLP EEQNFSGLRQRVQEMFPDRQINFSTSYNILTIAALVANSMDLAIIPSRFYN
LFSRCWPLEKLPFPPLNNEEQIDFSIHYNKFSLRDPILHGVIDVIRNAF"
complement(9578..10153)
/gene="lipB"
/note="synonym: b0630"
CDS complement(9578..10153)
/gene="lipB"
/function="enzyme; Biosynthesis of cofactors, carriers:
Lipoate"
/note="f191; 99 pct identical to LIPB_ECOLI SW: P30976"
/codon_start=1
/transl_table=11
/product="protein of lipoate biosynthesis"
/protein_id="AAC73731.1"
/db_xref="GI:1786848"
/translation="MHEFTDTRDDSTLDEIWLVEHYPVFTQGQAGKAEHILMPGDIPV
IQSDRGGQVTYHGPQQVMYVLLNLKRRKLGVRELVTLLLEQTVVNTLAE LGIEAHPRA
DAPGVYVGEKKICSLGLRIRRGCSFHGLALNVNMDLSPFLRINPCGYAGMEMAKISQW
KPEATTNNIAPRLLENILALLNPNDFEYITA"

BASE COUNT 2773 a 2298 c 2547 g 2646 t

ORIGIN

```
1 tttaatgatt ttaagttttt taattaatgt aattacgaaa tgactcgcag gttaagtga
61 tttaattgat ttaatgaata aaatttgcca cgatcataat taatatctat gtattttgat
121 tcaacatttt aattacatcc gtcaaagagg ctcgggacaa cccgcaagga aaacaatggt
181 gcagcttaac gagaataaac agtttgcat tttccaaaga ctggcatttc cgctgcgtat
241 ctttttgctg attctggtgt tctcaatatt tgtcattgca gccctggcgc aatattttac
301 ggccagtttt gaggactatt taacgcttca tgtacgcgac atggcaatga atcaggcgaa
361 aattattgcc tccaatgaca gtgtcatctc tgcggtgaaa acgcgtgact acaaacggct
421 ggcgaccatc gctaacaaat tacaagaga taccgatttt gattatgtgg tgattgggga
481 ccggcactcg atccgccttt accatcctaa tccggagaaa attggttatc ctatgcagtt
541 caccaaacag ggcgcgctgg agaaagggga gagctacttc attaccgga aagggtaaat
601 ggggatggcg atgcgcgcca aaacgccaat ctttgatgac gatggaaaag tcatcgcgct
661 ggtgtcgatt ggctacctgg tgagtaaaat cgatagctgg cgggctgagt ttttattacc
721 gatggcaggt gtgtttgtcg tgctgttagg gattctgatg ttgctgtcgt ggttcctggc
781 cgcgcatatc cgtcggcaga tgatgggcat ggagccaaag caaatcgcac gcgtgggtccg
841 tcagcaagag gcgctgttta gttcggttta tgaagggctg attgcggtgg atccgcatgg
901 ttacattacc gccatcaatc gtaacgcaag aaagatgctg gggctgagtt ccccgagacg
961 gcaatggttg ggtaaaccce ttggtgaagt ggtcaggccc gccgatttct ttaccgaaca
1021 gattgatgaa aaacgtcagg atgtggtggc gaactttaac ggtctgagcg ttattgcca
1081 ccgggaagct attcggtcag gtgatgatt gctgggggcc attatcagct ttcgtagtaa
1141 agacgaaatt tccaccctca atgcgcaat gacgcaaata aaacaatacg ttgagacctt
1201 tcgtacattg cgacacgagc atctcaattg gatgtcgacg ctcaatggtc tgttgcatat
1261 gaaagagtat gatcgctgac tggcgatggg gcagggggag tctcaggccc agcaacagct
1321 tattgacagc ctgcgcgagg cgtttgccga tcgccagggt gcggggctgc tttttggtaa
1381 agtgcagcgc gcccgggaac tggggctaaa aatgatcatt gtccccggtg gccagctttc
1441 gcaactgccg ccaggactgg atagcaccga gtttgacagc attgtgggca atttacttga
1501 taacgccttc gaagccagcc tgcgtagcga tgaaggaaac aagatcggtg aattattcct
1561 cagcgatgaa ggcgatgatg tgggtgattga agtcgccgat cagggtgcg gcgttcacga
1621 gtctctacga gacaaaatat ttgagcaggg ggtcagtagc cgtgctgacg agcccgtga
1681 acatggcatt ggggtgtact tgattgccag ctacgtaacg cgctgcggtg gtgttatcac
1741 tctcgaagat aatgatccct gcggtacctt attttcaatc tatattccga aagtgaacc
1801 taatgacagc tccattaacc ctattgatcg ttgaggacga aacgccgctg gcagagatgc
1861 atgcggaata tattcgtcac attcccggat tcagtcagat attactggcg ggaaatctgg
1921 cgcaggcccg aatgatgac gagcggttta agccggggct aatcttgctc gataactatc
1981 ttctgacgg tagagggatt aatttactgc atgaactggt gcaggcgcat tatcccgcg
2041 acgtggtggt taccactgca gccagcgata tggaaacggt gtctgaagcc gtacgttggt
```

```
2101 gtgtatttga ttatctcatt aagcccattg cctatgaacg gctggggcaa acgctaaccc
2161 gtttccgcca gcgtaaacad atgctggaaa gtattgatag cgccagccag aagcaaattg
2221 atgagatgtt taatgcttat gctcgcggtg aacctaagga cgagctaccg accggcattg
2281 accccttgac gctaaacgcg gtgcgaaaac tgtttaaaga gcctgggtgt caacatacgg
2341 cagaaacggg ggcgcaggca ctgaccatca gccgcaccac tgccaggcgt tatcttgaat
2401 attgcgccag ccgccatctg attattgctg aaattgttca cggcaaagt ggagaccac
2461 aacgcataata ccacagtggg tgacatgata aagccggagg aaacttgcc cggcattac
2521 tatttacttg cctgtgaccg ctgctgcccgt tcctggcacc atcagctctg tagcaacgat
2581 aacaatcacc aaaccaacaa gcaccgggtac cgagggtgctg tttacgactt caaacggcga
2641 gatcttcgcc atcccggcaa ccgcaacgac tacgccagaa acgggcgaaa gggtagggcc
2701 aaggtttgac gcctgcagca tcgggatagt caaatacgcc ggggttaatgc cggaagagtg
2761 cgccagtttc gggatcatct caacaaacgc ataaaacggc gcattgcctg aaccggctcg
2821 gactgccgcc agcattgtca gaatcaccaa taccagcatc aggatgatac tcgccgaacc
2881 aaacgaggta gcgatagaaa tcagactttg aataaagccg atggtgctaa gcccctgagc
2941 gaatacccca gcggcaacca gcagcatcac cacgttagca aacgcatctg ccattcccgc
3001 ataagccact tccagaccag agaaaacttt ctgggtatta aagctgcgga ggaactccag
3061 aatggaggca atcagcatat aaatcaccag aatagtgatg atgtgttaatt gcggaccca
3121 tttaccgtca aaaatcagta cccgatgat cgccgtgaac ggcaaatgg cataaacgc
3181 aggagcagtg gtgggtgattt cactgacatc taacatttca tgagagatgt gctctttttt
3241 atccagataa cgttgccaga agaagtgggc gatcgccatg ccgataattg cagcaattga
3301 gataggcagc gtcgttttga aggcgaagtc aatcagcgac atttcggaag cttgcgccgc
3361 cagcaccaca tcccctgaag tcggtgcgag aataatcgcc gccggggagg caaaaatggc
3421 agcagctgcg ccacgactga taccaacggt taccatcacc ggaaataggg ttgccatcag
3481 caaacacccc agaccggttg cggaagagac ggccagagac atcagacagg cgacaaaata
3541 ggcggcaatc atcagcaggt aaggggagtt aataactgac aatggttttg acgccagctt
3601 gaccaccata tcattcgcgc cgatatgggt catgtaagcg gcaaatccac acagcatcat
3661 aatcatcatg ccgaggtcgc cgccgcggct cattagtaat attttaacgt attcaacgat
3721 atccgtggcg ctgtagcctg ttgaagcctg gctggacggt aacactttgt gcccataat
3781 ggcactgata atcagcaata acaggccacc gacaaataac acaccagtgg cggaataccc
3841 tttaatgatg tagcgagcta ccccacaat aaccacaacc ccaataagga gctcaatgaa
3901 tgtcagcata atttttcctg tctccaggcc ccaaagtaaa taataaaaaa ttcctaaagc
3961 ttaaggaaaa aatatgccca ataaattggc gatgaatgct gattaaaatc aagaaaaact
4021 gccattaaga cattgaagtt gctgttttta tacatagata acaaacgcga tatttagcgg
4081 ttggcaatta tcacataata atgtgtgcaa attgatgggt gattgatgca aatttgttaa
4141 ttgcaatggg gttagctgcg gctgtgcgct caaaaaataa tctaatatga gcatagggtg
4201 acgatataata ttatatactt cgtaagatg attgttgtat ctcgttaaaa aataaaataa
4261 tttttccttg attgcatttt gtcatacaaa aagacttggg ttttcttttt tgactattcc
4321 catcgagaaa aacgacgcac catctttaat cgatgcgcgg aaatatatta cttgaacaag
4381 cggaaataaaa tagagcagct attcagatta ttctttatgt tgggtctatt aaggttatgt
4441 taattgtagc tttgctatgc tagtagtaga tttttgataa atgttttatg gtcacaaatg
4501 aacgtgagta aatatgtcgc tatcttttcc tttgttttta ttcagttaat cagcgttggg
4561 aaagtttttg ctaacgcaga tgagtggatg acaacgttta gagaaaatat tgcacaaacc
4621 tggcaacagc ctgaacatta tgatttatat attcctgcca tcacctggca tgcactgttc
4681 gcttacgaca aagaaaaaac cgatcgctat aacgagcgac cgtgggggtg cgtttttggc
4741 ctgtcgcggt gggatgaaaa aggaaactgg catggcctgt atgccatggc atttaaggac
4801 tcgtggaaca aatgggaacc gattgccgga tacggatggg aaagtacctg gcgaccgctg
4861 gcggatgaaa attttcattt aggtctggga ttcaccgctg gcgtaacggc acgcgataac
4921 tgggaattaca tccctctccc ggttctactg ccattggcct ccgtgggtta tggcccagtg
4981 acttttcaga tgacctacat tccgggtacc tacaacaatg gcaatgtgta ctttgctggtg
5041 atgcgctttc agttttgaga caaatgaagt ttagtaact tctttaaaat caatagctaa
5101 aataagtaac atcaaaaata acgcgacttt tatcactttt tagtaaagtt acactggaca
5161 aagcgtacca caattgggtg actggttaacc gacacagcat ttgtgtctat tttcatgta
5221 aaggtaatth tgatgtctaa gattaaaggt aacgttaagt ggtttaatga gtccaaagga
5281 ttcggtttca ttactccgga agacggcagc aaagacgtgt tcgtacactt ctctgcaatc
5341 cagactaatg gttttaaaac tcttgctgaa ggtcagcgcg tagagttcga aatcactaac
5401 ggtgccaaaag gcccttctgc tgcaaacgta atcgctctgt aagatacgtc agcaagaatt
5461 caaaacccgc ttaatcagcg ggtttttttt ggtcttttagt gtgcggttga ggccgaaaac
5521 agccagaatg ccagtgcggt catggcaaaa gacccagaaa ggttgacgaa aacgttcagt
5581 aatgcccgac caaagcggcc ctcttgtaac aaaaacacca cttctgccga aaatgttgag
5641 aaggttggtta gaccgccaca aaatccggtg gtgattaata ctttccacac tggatcaatg
```

```
5701 ttcgtcatcc tgcgaacca tgcgaatcct attcctatga tgaatgcccc aatcaggttt
5761 gctgtcagcg tccccaacgg aatcgcctga tgcagtgggt taaatcgcat acttaacagc
5821 catctcgcca cgcttcccggt accaccgcca ataaaaactg ctaaaagaag ttgtaacact
5881 gcaaaatcct gctatttgat ttgtatgagt gataagtgtg acgccgaata atcgtcgttg
5941 gcgaatttta cgactctgac aggaggtggc aatgctgggt gccgcaggac agtttgctgt
6001 tacatctgtg tgggaaaaga acgctgagat ttgtgcctcg ttgatggcgc agggcgcgga
6061 aaacgacgca tcgctgtttg ccctgccgga agcattgctg gcgcgcgatg atcatgatgc
6121 agatctatcg gttaaatacag cacagctgct ggaaggcgaa ttctcgggac ttacggcgga
6181 gaaagtaaac gtaacatgat gacgacaatt ctgacgattc atgttccttc aacgcggggg
6241 cgcgcagcga atatgctggg ggcacttcag gcaggaaaca tcgtcgcccc ttatgccaaa
6301 ctgcatctct atgatgcatt tgccattcag gaatcacgcc gtgttgatgc tggtaatgaa
6361 atcgctccgt tactggaggt ggaaggggat aaggtcggtc tgatgacctg ttatgactta
6421 cgctttccag agctggcgct ggcacaggca ttacagggag ctgaaatcct ggtacttcct
6481 gccgcctggg ttcgcggggc gctcaaagag catcactggt caacgttgct tgccgctcgt
6541 gcgctggata ccacctgtta tatggtggcg gcgggggagt gcgggaacaa aaatatcggt
6601 caaagccgga ttatagatcc ctttgccgtc accattgcgg cagcgtcaga aatgcctgca
6661 ctcatatagg cggaagtgcg gcccgaaacgt gtgcgtcagg tgcgcgcgca actgcccgtc
6721 ttaaacaacc gtgcgttttg gccgcgcgaa ttattatgat gtttttttac tcggcgcttg
6781 attcaccttg ttacagattg ctattgtgtg cgcgcgtcga atgaccgtta atattctctg
6841 gtttttaagg cgcgttctgt tgccggttat atgtcaagaa ggtatctatg ggtgagatta
6901 gtattaccaa actgctggta gttgcggcgc tggtcgttct gctgtttggg actaagaagt
6961 tacgtacgct gggcggagac cttggagcgg ccattaaagg gttcaagaag gcgatgaatg
7021 atgacgatgc tgcggcgaaa aaaggcgag acgttgatct tcaggctgaa aagctctctc
7081 ataaagagtg acgtggcgag caggacgctc cctcaatac ttgttcgata caaaaaaccc
7141 gcttcaaaaa gcgggttttt tatcagacag atgtaagtaa ttattacagg attacttaac
7201 ttccatccct ttgccttgca aatcggcgtg gtaagaagag cggacaaacg gaccgcatgc
7261 agcatgggta aagcccatcg ccagcgcttc ggctttcatt tcgtcgaact catccgggct
7321 aacgtaacgt tgaaccggca ggtgatggcg gcttggtgct aaatattgcc ccagcgtaaa
7381 catcgtcaca ccatgacggc gcaggctcgc cattacctca ataatttctt cattgggttc
7441 acccagtcac accatcagac cagacttggt cgggatttcc ggatgcgctt ctttaaagcg
7501 ttccagcagc ttcagcgacc agttgtaatc tgcaccaggc cgtacctgac ggtaaatacg
7561 cggtagcttt tccaggttat ggttgaacac atctggtggc gttgcagtca gaatatccag
7621 agcacgatcc atacgaccgc ggaaatccgg caccagagtt tcaattttga tttgcgggct
7681 tttttccoga atggcagtaa tgcaatccgc aaagtgcctg gcaccgccat cgcgcaggtc
7741 atcacgggta acggaggtga taaccacata acgcagcgcc atatcggcaa tggctctgcg
7801 cagtttcaact ggttcatttg catcaggagc taccgggcca ccgtgggcaa catcacagaa
7861 cggacaacgg cgggtacaaa tagcgccgag gatcataaac gttgctgtgc cgtggttgaa
7921 gcattccgcc aggttagggc aggaggtctc ctgcgagaca gaatgcaggc catttttgcg
7981 cattgcggtt ttgatgccct ggatacgtgt agagtccgct ggaagcttga ttttcatcca
8041 ttccggcttg cgcagcaggg cttcgcgctc tgttgccacg tttttaaccg ggataagggc
8101 catcttatcg gcatcgcggt atttaacacc gcgttccatc acaatgggtt tactcatagc
8161 gtgcgtgttc cagttgcgaa ttacgaagga aagcgtttca attcaataga ttgttgtaat
8221 tatcaactat tttgaatta acgactggca gtatatcatt gaaacggacc tgaagcagc
8281 caaagcggtc gcgaaaatgt aaaattgttg ttggattgtg ccattttatc gttctgcgct
8341 gtgatctggc tgtaacaata ctttttcaaa atgcattacg gataacatct attactcgt
8401 gcaggatggg atcacgcagg ctgaatttgt tgtagtggat agagaaatct atttgctcct
8461 catttaagga cggaaaaggc agtttttcca gcggccagca gcggctaaac aggttataaaa
8521 aacggctggg aataatcgct aacatgtcac tgttggaac cagtgcagcg attgtcaaaa
8581 tgttgtagct ggtgaaattg atttgccggt ccggaacat ttcttgaact ctctgccgca
8641 gaccgctgaa attttgctct tccggtaaca gaagtacatg cgcagcgttg tcgatagtct
8701 ctctgtcatc ttctaaagag agtagtggtt ttccctcgcg gcaaattaac accatatgt
8761 cggatgaacag aacatgatgt tgcaccgtac gattggtgca aaacatgtta tcgatgatga
8821 gatcggtttg aaactgactg agttgggttt ccgcgtcgct gatgggcggg ttgcgcagta
8881 aaagctgcgg atagtgaagt ttaatcgccc gatagatgac aggaaggacc agtgctccga
8941 ccgagggagt ggtcgcaata gttatcggtt gttgcttatc atagcttctc tcgatatcca
9001 gcgcgccaag aatggactca aggcctgac tgatatactc atgtagatgc atcgcaaatg
9061 cggtaggagt gacaccctgg cctttgcgaa taacaatgg gtcagggaat ataacgcgca
9121 gttttctgaat agactgactg attgccgagg ggtcagatt aagcactttc gctgcattaa
9181 cgatcccttt atgtacatat acagcttcaa aaatagtcag aaggttaaga tcaatatctc
9241 gtaaggttgt aaatatattg ggtttacctt ctgacgactt tctgcttaaa cagggttcaa
```

9301 tttgattatt actatccacg cacttactcc aatttttattc atggaaaaat aatattttaa
9361 aaattacaat aatcttatgt ctaattggaa cggaacgctt ttgctcacca taatcaacta
9421 tttcaatagg ttaatcaatg gggtaggtta cggagcaagg cgtggcaccg catttctgca
9481 agtgataaag aaaatataca aacgcgaagg agatgtaaag cattagcaga tattatgagt
9541 aatgacccag tgtaaattgg gccattgatg tatggaatta agcggtaata tattcgaagt
9601 ccggattggt tagtagcgct aaaatatttt ccagttaaagc tggagcaata ttattagtcg
9661 tcgcttcggg tttccattgt gatatttttag ccattttccat tccggcatac ccacaaggat
9721 taatacgtaa aaatggtgaa agatccatat tgacgtttta tgccagaccg tggaatgaac
9781 aaccgcgctc aatacgtaaa ccagtgagc aaattttctt tcccccaaca tagacacctg
9841 gcgcgctcagc ccgaggatgc gcttctatac ccagttcagc cagggtattc accactgttt
9901 gctcaagcaa ggtcaccagt tcacgcacac cgagtttacg gcgtttcagg ttaagcaaca
9961 catacatcac ctgttgcccc ggcccgatg aagtcacctg cccaccgca tcgctctgga
10021 tcaccggaat atcaccggc attaaaatgt gctccgcttt tctgcctga ccttgggtga
10081 ataccggata gtgctcgacc agccagattt catcaagggt actatcatcg cgggtatcgg
10141 tgaattcatg catagcctgg gagattggct cgtaaggctg aagaccgagc tggcggacaa
10201 gaattttatc ctgatacaaa acggcatctc cgtggagaat gagtaaaagt ggggggaaaag
10261 tata

//

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

May 20 2003 11:20:12